



Research papers

Change-signal impacts in downscaled data and its influence on hydroclimate projections

N.J. Potter^{a,*}, M. Ekström^{a,1}, F.H.S. Chiew^a, L. Zhang^a, G. Fu^b^a CSIRO Land and Water, GPO Box 1700, Canberra, Australian Capital Territory, 2601, Australia^b CSIRO Land and Water, Private Bag 5, Wembley, Western Australia, 6913, Australia

ARTICLE INFO

This manuscript was handled by Marco borga, Editor-in-Chief, with the assistance of Ashish sharma, Associate Editor

Keywords:

Downscaling
Hydroclimate projections
Climate change
Runoff
Victoria, Australia

ABSTRACT

Downscaling is the process by which output from global climate models is translated to finer resolution regional scale projections often used in impact studies. Many fundamentally different techniques can be used, each with different capabilities of resolving or representing sub-gridscale processes. The different formulations can lead to variations in the downscaled output with consequential impacts in the interpretation of future change. Here, future runoff is estimated for six catchments in the Australian state of Victoria using five different regional rainfall projection products. We investigate how differences in rainfall input manifest in a selection of regionally important hydroclimate metrics. Overall, annual runoff is projected to decline under most methods, but seasonal changes are more uncertain reflecting differences in the rainfall change signal for different downscaled products. Whilst change in flow metrics are mostly consistent with rainfall change factors, changes in low flow (e.g. 7-day minimum flow) show considerable uncertainty, especially for drier, ephemeral catchments. Results from empirical (simple) scaling of climate observations generally lie within the range of more complex downscaling methods. However, empirical scaling is unable to provide meaningful information on spatial heterogeneity in the change signals, as well as for several metrics of rainfall and runoff. Other downscaling methods can potentially provide information on these, but the large uncertainty remains a problem, as well as our currently poor understanding of method-related biases.

1. Introduction

Planning for the future is an integral part of managing water resources (Poff et al., 2016), even more so in regions with large natural variability and periodically scarce supply. Such are the conditions in southeast Australia, where long-term trends show an ongoing decline of cool season (April to October) rainfall since the early 1990s (Hope et al., 2016). Since the cool season is the ‘filling season’ for regional reservoirs in southeast Australia, there is concern among water users that the observed decline may continue into future decades (Potter and Chiew, 2009, 2011; Potter et al., 2011). The current national climate projections for the region suggest high confidence in drying in southern Australia, particularly in spring (September to November) as a response to shifts in westerlies and strengthening of the subtropical ridge (CSIRO and Bureau of Meteorology, 2015). These findings further enhance concern for future water supply in the region.

To provide meaningful information for water resource impact assessments, the coarse spatial signal from global climate models (GCMs) needs to be downscaled to finer resolution scales (i.e. point or

catchment scale). Because downscaling methods vary in technical formulation and complexity, for some variables and regions, downscaling methods can produce very different results. Many downscaling comparisons have been conducted to demonstrate and understand the plausible range of projections in downscaling ensembles (see, e.g. Fowler et al., 2007; Chiew et al., 2010; Chen et al., 2011; Frost et al., 2011; Grose et al., 2015; Sunyer et al. 2015).

The State Government of Victoria (Australia) provides guidance about future change to runoff for water corporations, catchment management authorities and industries to manage and plan for long-term risks to water supply (DELWP, 2016). Guidance for a 50-year time horizon is contained in the State’s Water Supply and Demand (WSD) plans, a document that is revised every 5 years. The most recent hydroclimate projections supporting the WSD planning considered both GCM projected change and observed decadal variability. The GCM based projections used climate change information from 42 models from the fifth Climate Model Intercomparison Project (CMIP5, Taylor et al., 2012), downscaled to the region using a change factor approach (Potter et al., 2016). However, as recently shown by Ekström et al

* Corresponding author.

E-mail address: Nick.Potter@csiro.au (N.J. Potter).¹ Present address: Cardiff University, Wales, CF10 3AT, UK

Table 1

Downscaling methods available for this study. © 2018 CSIRO. All Rights Reserved.

Downscaling method	Emissions scenario	Historical	Future	Ensemble size	Spatial resolution	Reference
Empirical scaling	RCP8.5 ¹	1986–2005	2060–2079	42	0.05° × 0.05°	Potter et al. (2016)
Statistical analogues (SDM)	RCP8.5 ¹	1986–2005	2060–2079	22	0.05° × 0.05°	Timbal et al. (2009, 2011); Teng et al. (2012a)
Non-homogenous Hidden Markov Model (NHMM)	RCP8.5 ¹	1986–2005	2060–2079	19	Catchment-based	Charles et al. (1999); Fu et al. (2013a)
Conformal-Cubic Atmospheric Model (CCAM)	RCP8.5 ¹	1986–2005	2056–2075	6	0.5° × 0.5°	McGregor and Dix (2008)
Weather Research and Forecasting model (WRF)	SRESA2 ²	1990–2009	2060–2079	4 × 3 ³	0.1° × 0.1°	Evans et al. (2014a)

¹ RCP8.5 is the high emission Representative Concentration Pathway (RCP) (van Vuuren et al., 2011) used by the latest generation of GCMs in the Coupled Model Intercomparison Project Phase 5 (CMIP5) (Taylor et al., 2012).

² SRES A2 is the high emission scenario in the Special Report on Emissions Scenarios (SRES) (Nakicenovic et al., 2000) used by the previous generation of GCMs (CMIP3).

³ 4 GCMs with 3 different physics scheme configurations of WRF.

(2016), there is emerging evidence of a regional change signal in available downscaling products that is not reflected in GCM-based projections for southeast Australia. The nature of this change signal differs between different downscaling products making it difficult for users of these datasets to assess their relative value in comparison to the GCM based projections.

If using climate change projections for policy guidance, researchers need to carefully consider the choice of methods used and sources of information. Of primary importance is the representation of key sources of uncertainty in global climate change information, and secondly, ensuring that the method or methods used to derive a regional resolved projection have the capacity to translate the aspect of change relevant to the application (e.g. to spell duration or extremes). The key sources of uncertainty are found in: (1) the emission scenarios used for the global climate modelling (uncertainties about the species and rate of greenhouse gas (GHG) emissions); (2) the models used to simulate the climate response to the changing GHG emissions (aspects of model structure and parameter schemes that influence the models ability to simulate the climate); and (3) natural climate variability (the ability of the model to simulate the full range variability in the climate not arising from emission forcing). To attempt to represent these sources of uncertainty, researchers can consider GCM output using a range of emission scenarios (such as the Representative Concentration Pathways (RCPs) (van Vuuren et al., 2011), and consider the use of a large number of GCMs (to sample model ability to simulate forced and internal climate variability).

Each additional manipulation of data adds some level of uncertainty to the outcomes, such as the downscaling step and the impact modelling itself. For example, many hydrological models are conceptual models, relying on a robust calibration to provide good estimates of streamflow. Under changing climate conditions these assumptions may no longer hold. This challenge is often referred to as the non-stationarity problem, discussed at length by several authors (Milly et al., 2008; Potter et al., 2013; Saft et al., 2016; Chiew et al., 2017). An additional concern for projections is the choice of baseline period. This is the choice of time window for which the future data is compared and represents the future climate (see e.g., Timbal et al., 2016; Potter et al., 2016). Further details on the production of climate change datasets and the impact on policy guidance is detailed in Harris et al. (2014) and Ekström et al (2016).

In this study, we examine rainfall from the five different downscaled datasets that are currently available for six representative catchments in the State of Victoria. The aim is to quantify differences in rainfall projections by different available downscaled climate change datasets, and further examine how such differences manifest in metrics derived from streamflow modelled using the different downscaled datasets. The ensemble of datasets is an ensemble of opportunity, representing all downscaled datasets available for the region in 2017. Because of the opportunistic nature of this ensemble, there is some variation in the use

of emission scenario, CMIP archive and slight difference in time horizons. However, all selected projections have the common purpose of illustrating regional climate change for a far-future time horizon under a high-emission scenario, and the ensemble represents the current information available to researchers, planners and purveyors of climate change information.

Section 2 outlines the downscaling methods, climate data and hydrological modelling used in this study. Section 3 presents changes in mean annual and seasonal rainfall across Victoria, annual and seasonal runoff for six study catchments, and changes in key rainfall and runoff metrics, as well as the differing spatial resolution of the different downscaling methods and the potential for providing within-catchment climate information. Section 4 discusses limitations of the study and directions for future research, and Section 5 gives conclusions and recommendations.

2. Data and methods

2.1. Downscaled rainfall datasets

There are many fundamentally different approaches used to achieve a finer resolved climate projection. It is well established that different methods can influence what aspects of the GCM 'change signal' are translated to the finer resolved downscaled data set (e.g. Fowler et al., 2007; Chen et al., 2011; Ekström et al., 2015). Briefly, downscaling methods can be categorised as: (1) empirical downscaling or 'change-factor methods', which apply projected changes in the distribution of GCM rainfall to observed (point or catchment-scale) climate data; (2) statistical downscaling methods, which use statistical relationships based on observed data and apply these on GCM output to derive a local variable estimate; and (3) dynamical downscaling methods, typically involving regional climate models (RCMs) that simulate physical processes at a finer spatial scale using boundary conditions from a host GCM.

This study considers all currently available downscaled datasets in south-eastern Australia with output that is readily available for hydrological modelling (Table 1). The datasets are produced by different research initiatives across the region: the Victorian Climate Initiative (Hope et al., 2016) and its predecessor, the South Eastern Australian Climate Initiative (CSIRO, 2012), the New South Wales and Australian Capital Territory Regional Climate Modelling (NARClIM) project (Evans et al., 2014a), and two datasets created to support the national climate projections (CSIRO and Bureau of Meteorology, 2015). The methods used to create these datasets are fundamentally different ranging from empirical scaling to dynamical downscaling. Furthermore, there are some differences in the projected time horizon, GCM selection and emission scenario used (Table 1). Table 2 provides a full list of the GCMs downscaled by each method. Mean annual regional rainfall for

Table 2

Global climate models downscaled by each method. Note that each CMIP3 GCM downscaled by WRF is downscaled with an independent regional climate physics scheme, resulting in the ensemble size of 12 reported in Table 1. © 2018 CSIRO. All Rights Reserved.

	Empirical scaling	SDM	NHMM	CCAM	WRF
<i>CMIP3</i>					
CGCM3					X
CSIRO-Mk3.0					X
ECHAM5					X
MIROC3.2					X
<i>CMIP5</i>					
ACCESS1.0	X	X	X	X	
ACCESS1.3	X	X	X		
BCC-CSM1.1	X				
BCC-CSM1.1(m)	X	X	X		
BNU-ESM	X	X			
CanESM2	X	X	X		
CCSM4	X	X		X	
CESM1 (BGC)	X				
CESM1 (CAM5)	X				
CESM1 (WACCM)	X				
CMCC-CESM	X				
CMCC-CM	X				
CMCC-CMS	X	X	X		
CNRM-CM5	X	X	X	X	
CSIRO-Mk3.6.0	X	X	X		
EC-EARTH	X				
FGOALS-g2	X				
FGOALS-s2	X		X		
FIO-ESM	X				
GFDL-CM3	X			X	
GFDL-ESM2G	X	X	X		
GFDL-ESM2M	X	X	X		
GISS-E2-H	X				
GISS-E2-H-CC	X				
GISS-E2-R	X				
GISS-E2-R-CC	X				
HadGEM2-AO	X				
HadGEM2-CC	X	X			
HadGEM2-ES	X				
INM-CM4	X		X		
IPSL-CM5A-LR	X	X	X		
IPSL-CM5A-MR	X	X			
IPSL-CM5B-LR	X	X	X		
MIROC-ESM	X	X	X		
MIROC-ESM-CHEM	X	X			
MIROC5	X	X	X		
MPI-ESM-LR	X	X	X	X	
MPI-ESM-MR	X	X	X		
MRI-CGCM3	X	X	X		
MRI-ESM1	X				
NorESM1-M	X	X	X	X	
NorESM1-ME	X				

twenty-year climate baselines are within $\pm 2\%$ of the 20th century annual average (Potter et al., 2016), and the historical time periods are consistent with current recommendations for choosing baseline climate periods for southeast Australia (Timbal et al., 2016). While acknowledging differences among datasets, we use these datasets in a comparative study to demonstrate the range in streamflow projections for a far-future period under a high-emission scenario. Note that for consistency we use historical potential evapotranspiration for the hydrological modelling, as described in Section 2.2. Each dataset is described below.

It is important to note that precipitation data from the two dynamical downscaling methods (WRF and CCAM) were bias corrected (as described in Sections 2.1.4 and 2.1.5). We did not bias correct the statistical methods (NHMM and SDM) since these methods have already been constrained to observed rainfall gauges in the region. That is, we consider the data processing methods used by the statistical methods sufficient for hydrological modelling without further post-processing. In contrast, the dynamical methods are designed to represent the regional

climate dynamics holistically. As such, precipitation is not explicitly constrained to observations with the result that dynamical methods typically produce biases in rainfall occurrences and amounts (Teutschbein and Seibert, 2010), necessitating bias correction before hydrological modelling.

2.1.1. Empirical scaling

Empirical scaling has been recommended for use in regional water availability studies (Frost et al., 2011) and underpins recent runoff projections for south-eastern Australia (CSIRO, 2008; Chiew et al., 2009; Post et al., 2012) and Victoria (Potter et al., 2016). The premise of empirical scaling is that climate information from GCMs (particularly rainfall) is not directly usable in hydrological models since the distribution of GCM rainfall is different from the observed catchment-scale daily rainfall distribution. However, empirical scaling assumes that GCMs have sufficient skill to reasonably model future changes in the annual and seasonal rainfall averages. These annual and seasonal change factors are then extracted from the GCM historical and future runs and applied to observed (catchment or station) rainfall and potential evapotranspiration time series. Empirical scaling therefore relies directly on the observed climate time series and does not consider potential changes in rainfall occurrences and sequences and other finer-scale temporal variability.

The observational datasets used here are the gridded daily rainfall and temperature observations of the Australian Water Availability Project (AWAP, Jones et al., 2009). The same gridded datasets were used for empirical scaling to inform the most recent production of regional streamflow projections for Victoria (Potter et al., 2016). More details on these data can be found in Section 2.2. Here, scaling uses seasonal and annual scaling factors from 42 CMIP5 GCMs. Annual and seasonal means are calculated for the historical and future time periods (Table 1) for each GCM. Seasonal change factors are used to scale daily rainfall in each season individually, and the annual change factor is used to scale the seasonally scaled rainfall time series to preserve the annual change in GCM rainfall.

2.1.2. The Statistical Downscaling Model

The Statistical Downscaling Model (SDM, Timbal et al., 2009, 2011) was one of two downscaling methods used to complement the otherwise GCM-based national climate projections by CSIRO and Bureau of Meteorology (2015), the second being the dynamical downscaling output by CCAM – see Section 2.1.4. The SDM method seeks meteorological analogues of future atmospheric variables from the historical data to develop future rainfall projections, relating observed rainfall and temperature to reanalysis atmospheric predictors (mean sea-level pressure, temperature, rainfall, humidity and wind speeds). These statistical relationships are then used with future GCM projections of atmospheric predictors (including rainfall) to produce future time series of rainfall and maximum temperature.

The SDM was applied to 22 CMIP5 models (see Table 2), which were those GCMs with the necessary output variables required for the method, to produce daily and monthly climate time series across Victoria and Australia for downscaled CMIP5 GCMs). Due to post-processing needs, the daily resolved SDM was not available for this analysis. Instead a monthly-resolution version of the dataset was used to create seasonal and annual scaling factors that were used to scale observed gridded rainfall in the same way as in empirical scaling (see Section 2.1.1).

2.1.3. Non-homogeneous Hidden Markov Model

Projections for selected Victorian catchments from the Non-homogeneous Hidden Markov Model (NHMM) statistical downscaling method (Hughes et al., 1999; Charles et al., 1999; Fu et al., 2013a, 2013b) are available for selected catchments. In NHMM, daily rainfall amounts are stochastically generated from a distribution, which is dependent on hidden weather states. The sequence of weather states is

determined by variability in large-scale atmospheric predictors. Here, NHMM is downscaled directly to catchments rather than points (i.e. grid cells), which was found to better reproduce the rainfall distribution (Fu et al., 2013b). NHMM can reproduce reasonably reliably the daily rainfall characteristics in terms of amounts and distributions. This allows the dataset to be used directly as input into the hydrological model. NHMM downscales 19 of the CMIP5 GCMs, which are those GCMs with the necessary output variables required for the method.

2.1.4. The Conformal Cubic Atmospheric Model

The Conformal Cubic Atmospheric Model (CCAM) (McGregor, 2005; McGregor and Dix, 2008) is a variable resolution atmospheric model that ingest the GCM change signal through bias corrected GCM sea surface temperatures (SSTs). It has a stretchable grid across the globe, but for this experiment, CCAM outputs globally on a regular 0.5° resolution. As part of the national projections, CCAM was run with SSTs from six CMIP5 GCMs deemed to have high skill in simulating the regional climate (see Table 3.3.1, CSIRO and Bureau of Meteorology, 2015). An assessment of CCAM by Grose et al. (2015) notes a stronger rainfall response to increased SSTs compared to its host GCMs. This stronger relationship with SSTs could be indicative of a positive bias in its rainfall projections.

Whereas the previous three downscaling methods produce rainfall time series at point/catchment scale, and are produced for direct use in hydrological modelling applications, CCAM output covers larger grid cells (Table 1). An analysis of the CCAM rainfall data showed that the daily time series had too few rainfall days and a highly skewed rainfall amount distribution (i.e. too many large rainfall amounts) compared to observed point-scale rainfall time series. As such, the CCAM daily rainfall time series were bias corrected (i.e. relating CCAM daily rainfall distribution in historical climatology simulations to observed daily rainfall distribution, and using this relationship to convert CCAM future daily rainfall to grid or catchment scale future daily rainfall) to observed daily rainfall aggregated to either CCAM grid cells (for grid scale results, Fig. 2) or catchments (for the catchment results). To adjust biases, the double-gamma distribution mapping bias correction method of Teng et al. (2015) was applied to the CCAM output. This method has previously been shown to provide a robust correction of biases in this geographical region (Teng et al., 2015).

2.1.5. The Weather Research and Forecasting model

The Weather Research and Forecasting (WRF) model is a mesoscale atmospheric model with many applications both in numerical weather prediction and climate projections (Skamarock and Klemp, 2008). In a climate downscaling setup, WRF reads in output from a GCM along its lateral and lower boundaries and simulates the climate on a finer resolution within those boundaries. WRF was used to generate an ensemble of regional climate model simulations for New South Wales (NSW) and the Australian Capital Territory (ACT) (NARClIM; NSW/ACT Regional Climate Modelling) (Evans et al., 2014a). The NARClIM projections use three different configurations of WRF in combination with four GCMs that were selected to represent GCM uncertainty based on their skill and independence (Evans et al., 2013). As evident from Table 1, the NARClIM projections differ from other datasets in that they use climate change information from the previous generation of GCMs, i.e. GCMs from the CMIP3 archive (Meehl et al., 2007) that follow the A2 Scenario of the Special Report on Emissions Scenarios (SRES) (Nakicenovic et al., 2000). Nevertheless, as stated earlier, this reflects a similar high emission scenario towards the end of the century. This study used the bias-corrected (to AWAP) daily rainfall series available through the NARClIM data portal (Evans and Argüeso, 2014b).

2.2. Runoff and catchment hydrological modelling

Rainfall-runoff modelling is used to explore the effects of rainfall changes on streamflow and water availability. Six catchments were chosen from a subset of available gauges with good streamflow data completeness to represent a range of Victoria's climatic and hydrological variability (Fig. 1). Mean annual climate and streamflow, calibration Nash-Sutcliffe Efficiency (NSE) and catchment area are shown in Table 3. The majority of runoff in Victoria occurs along the Great Dividing Range to the east, and as such the western catchments (i.e. catchments 406213, 406214 and 407215) have ephemeral hydrographs with no baseflow sustaining streamflow during the drier months, typically autumn (MAM). The eastern catchments are more representative of water-supply catchments for Victoria and are perennial streams. Brankeet Creek (405251) is normally a perennial stream, but became ephemeral during the Millennium Drought (1997–2009) largely due to decreased water tables and catchment storage during the prolonged and

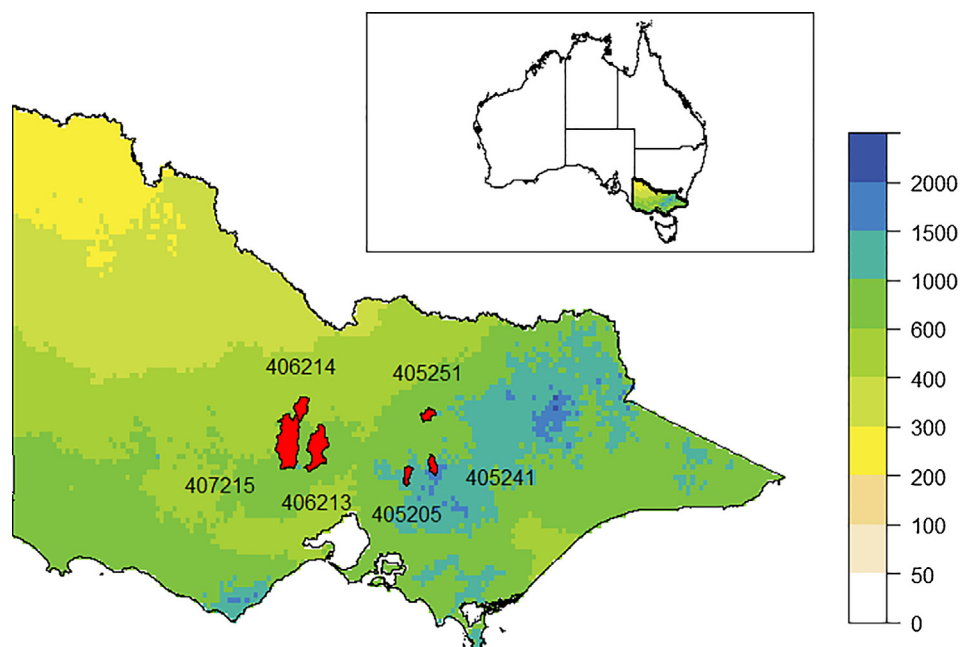


Fig. 1. Location of study catchments in Victoria. Inset shows the state of Victoria relative to Australia. The colour gradient over Victoria represents mean annual rainfall (mm) over 1975–2014. © 2018 CSIRO. All Rights Reserved.

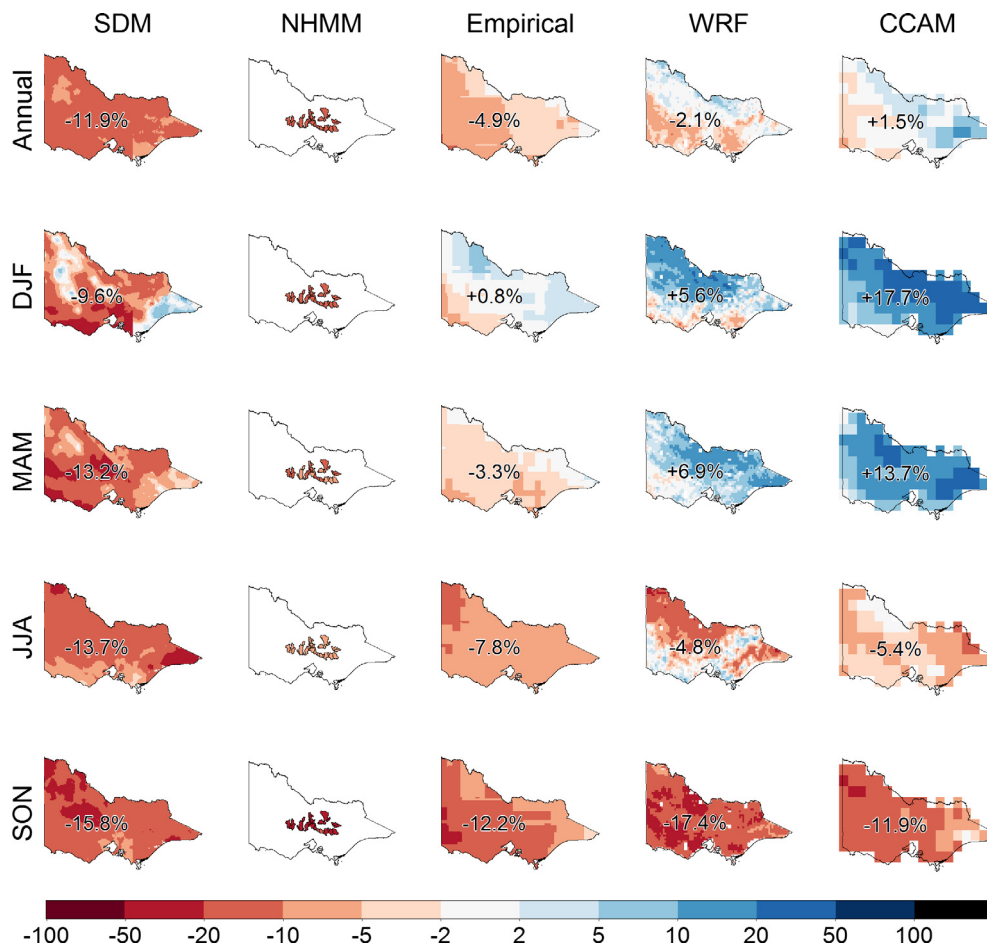


Fig. 2. Ensemble median annual and seasonal rainfall percentage changes around 2065. The downscaling methods are ordered by increasing change signal over Victoria (i.e. drier to wetter). WRF and CCAM rainfall is bias-corrected (see Sections 2.1.4 and 2.1.5). © 2018 CSIRO. All Rights Reserved.

severe drought during those years (CSIRO, 2012; Petheram et al., 2011).

Gridded daily climate data were sourced from the Bureau of Meteorology at $0.05^\circ \times 0.05^\circ$ (approximately $5 \text{ km} \times 5 \text{ km}$) resolution across the study region from AWAP. The $0.05^\circ \times 0.05^\circ$ gridded daily rainfall data (both unscaled and scaled to reflect future climate series) were aggregated to provide daily catchment rainfall time series. Potential evapotranspiration was estimated from solar radiation, vapour pressure, and minimum and maximum temperature using Morton's areal potential evapotranspiration (PET) formulation (Morton, 1983; Chiew and McMahon, 1991).

The conceptual rainfall-runoff model SIMHYD (Chiew et al., 2002) was calibrated to historical (1975–2014) streamflow data using the Shuffled Complex Evolution method (Duan et al., 1992) with the bias-

penalised NSE objective function of Viney et al. (2009). More details are provided by Potter et al. (2016). SIMHYD can satisfactorily reproduce the observed time series in these catchments, with daily NSE generally greater than 0.75, more than reliable for the purpose of this study.

Future runoff is modelled using the future rainfall time series for each downscaling method as described in section 2.1. In all cases, we use historical potential evapotranspiration (PET) time series to force SIMHYD for future projections of runoff, four key reasons being: (1) the different methods estimate PET differently (although we can estimate changes to PET by calculating PET from the changed climate variables); (2) compared to rainfall, PET is a secondary driver as the divergence among models of projected changes to PET is much smaller than projected changes to rainfall (Potter et al., 2016); (3) runoff is typically more sensitive to changes to rainfall than to PET in the medium-term

Table 3

Details of the study catchments. © 2018 CSIRO. All Rights Reserved.

Catchment ID	Catchment name	Mean annual streamflow (mm)					
		Baseline observed (mm)	Baseline modelled (mm)	Mean annual rainfall (mm)	Mean annual PET (mm)	Calibration Nash-Sutcliffe Efficiency	Catchment area (km ²)
405205	Murrindindi R. @ Murrindindi above Colwells	462	467	1270	1110	0.76	106
405241	Rubicon R. @ Rubicon	822	809	1410	1118	0.77	128
405251	Brankeet Ck @ Ancona	117	123	857	1220	0.67	123
406213	Campaspe R. @ Redesdale	103	106	761	1187	0.81	634
406214	Axe Ck @ Longlea	49	56	592	1262	0.75	235
407215	Loddon R. @ Newstead	69	70	680	1207	0.84	1028

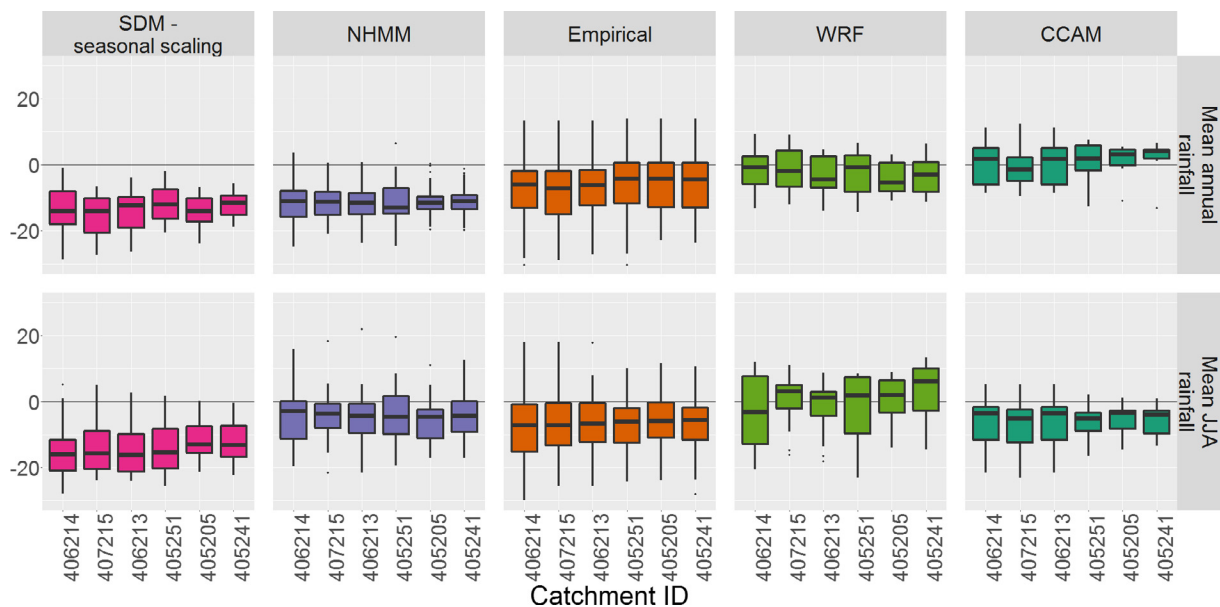


Fig. 3. Percentage changes to mean annual catchment rainfall (top row) and mean winter (JJA) rainfall (bottom row). As with later figures, the boxes here represent the interquartile range (25th and 75th percentile of results), with the median shown as the cross bar. The whiskers extend to 1.5 times the interquartile range, with outlying model results plotted as dots. © 2018 CSIRO. All Rights Reserved.

(Chiew, 2006; Potter et al., 2011); and (4) using the same historical PET allows us to explore changes in runoff from the projected changes in rainfall alone. We acknowledge that for other applications it may well be relevant to also use PET that includes a climate change signal.

2.3. Exploring metrics of extreme rainfall and flow

Whereas changes to annual and seasonal rainfall and runoff are relevant to understanding future broad-scale water availability, information on changes to these averages only is insufficient for many applications. For example, most dam inflows occur in response to high-flow events and, as such, water managers may be more interested in changes to high flows than average flows, as well as year-to-year variability in rainfall. Ecological resilience of a systems may depend on persistent low flow events (as well as periodic high flow events) to maintain hydrologic connectivity and volumes and water quality of floodplains, lakes and rivers. We explore the consistency and ability of different downscaling methods to model key metrics (see Section 3.4) based on rainfall and runoff that are deemed relevant to regional water managers, environmental managers and other modellers:

- Dry days and dry spells – proportion of days with rainfall less than 1 mm; and the mean duration of dry spell (defined as consecutive days with rainfall less than 1 mm).
- High rainfall – first percentile of daily rainfall (P1); and maximum 3-day rainfall total.
- Interannual rainfall variability – coefficient of variation (i.e. the standard deviation divided by the mean) of: annual rainfall totals; maximum annual rainfall; and probability of dry day (less than 1 mm) occurrence.
- Low flow – minimum 7-day runoff total; 90th percentile of daily runoff (Q90); and number of days below historical Q90_H.
- High flow – 5th percentile of daily runoff (Q5); number of days above historical Q5_H; and 1st percentile of daily runoff (Q1).

3. Results

3.1. Annual and seasonal median rainfall changes

Fig. 2 shows the median percentage change in annual and seasonal

rainfall from the GCM ensemble (which is different for each downscaling method, see Table 2). The numbers overlaid on the maps show the percentage change in rainfall regionally over Victoria. As noted in Section 2.1 (see Table 1), the NHMM data is downscaled to catchments, whereas the other methods are downscaled to grids of different resolutions.

The changes from empirical scaling can be interpreted as the direct changes from the raw GCM signals. The projected median change from empirical scaling (and therefore from the GCMs) is a 2–5% decrease in annual rainfall for eastern Victoria and a 5–10% decrease for western Victoria. Compared to the raw GCM change signal, SDM significantly enhances the projected drying. NHMM projections are slightly drier than the raw GCM signal. CCAM is considerably wetter than the raw signal, particularly for summer (December to February, DJF) and autumn (March to May, MAM), and in fact projects a median increase in annual rainfall of 1.5% averaged across Victoria. WRF projects a decrease in annual rainfall, approximately half of the decrease from the raw GCM signal. WRF projections of DJF and MAM rainfall are wetter than those from the GCMs. WRF also features a slight increase in winter (June to August, JJA) rainfall over the high-rainfall area of the Great Dividing Range (see Fig. 1), which has consequences for runoff production in the wetter catchments (as discussed later).

The projected rainfall decrease from the raw GCM signal occurs predominantly in winter (JJA) and spring (SON). Parts of south-eastern Australia, mostly north of Victoria, have projected increases in DJF rainfall (see CSIRO and Bureau of Meteorology, 2015; Potter et al., 2016), with the raw GCM signal in the north of Victoria partly reflecting this (Fig. 2). This increase in DJF rainfall in northern Victoria is also observed for WRF and CCAM. Recent research, however, contends that the DJF increase is due to GCMs that perform poorly according to several climate metrics relevant to Victoria (Hope et al., 2016; Grose et al., 2016). The increase in DJF is enhanced by WRF and especially CCAM, while SDM shows a significant decrease in DJF rainfall. Changes to MAM rainfall are most uncertain, with WRF and CCAM projecting increases but all other methods projecting decreases. The JJA signal is largely consistent except for slight increases due to WRF, and the SON signal is the most consistent with large projected decreases from all methods.

Fig. 3 shows the annual and JJA rainfall change factors extracted to the six selected catchments. Overall there is little difference among

catchments for each downscaling method except perhaps for WRF and CCAM. WRF has a slight tendency to project relatively more rainfall for the wetter catchments in JJA, and CCAM has the same tendency but for annual rainfall. Overall, however, the differences between downscaling methods are larger than the differences between catchments, which mirror the State-wide results in Fig. 2.

Fig. 3 also shows the range of results, and not just the median as in Fig. 2. The entire GCM ensemble spans a large range, which extends from projected increases at the wet end of the distribution to rainfall decreases at the dry end of the distribution (see also Potter et al., 2016; CSIRO and Bureau of Meteorology, 2015). This large range of results, featuring both increases and decreases, are also seen in NHMM, WRF and CCAM, but with a smaller range in the projections. SDM on the other hand projects decreases in future rainfall even from the wetter members of the downscaled GCM ensemble.

3.2. Bias in rainfall from different GCM model ensembles

The different downscaling methods each downscale a different set of host GCMs (see Table 2). The selection of a drier or wetter subset of GCMs would tend by itself to produce a drier or wetter change signal for the downscaling method. Fig. 4 compares the empirically scaled runoff change signal using the full set of 42 GCMs, the empirically scaled change factor using each method's GCM subset, and the downscaled change factor for SDM, NHMM and CCAM.

WRF was not considered for three reasons. Firstly, the 4 GCMs selected for the WRF modelling procedure were chosen to be fully representative of the GCM ensemble with an objective examination of the covariance in model errors (Evans et al., 2013). Secondly, the WRF data is based on CMIP3 and so we cannot directly compare the change signal to the restricted CMIP5 ensemble. Thirdly, the WRF GCM ensemble contains three different configurations of the model, which should each be considered as independent models, and again we are unable to use the CMIP5 ensemble to directly compare to this result.

The results indicate that for the statistical downscaling methods, the 22 GCMs used for SDM and, to a lesser extent, the 19 GCMs used for NHMM are drier than the full complement of 42 GCMs used for empirical scaling. This is true across the entire distribution (median and interquartile ranges). Thus, the GCM selection used by these methods results in a drier projection for both rainfall and runoff. However, SDM significantly enhances the decline in rainfall (and modelled runoff) projected by the GCMs. The change signal from NHMM is similar to that from the host GCMs, but with a smaller range. The six GCMs used for CCAM tend to be very representative of the full CMIP5 ensemble, showing a projected median drying of between 10% and 30% for the study catchments, but projections for annual runoff under CCAM largely show no change. This appears due to the quite different seasonal

change signals for rainfall from CCAM (Fig. 2) with runoff occurring from the projected increase in rainfall in DJF and MAM largely offsetting smaller runoff declines in JJA and SON (Fig. 5).

3.3. Projected changes to mean annual and seasonal runoff

Differences in the annual change signals for runoff in the six study catchments among downscaling methods closely follow the results for rainfall (Fig. 2). The percentage change signal for rainfall is enhanced in the percentage change signal for runoff, an effect often referred to as the rainfall elasticity (e.g. Chiew, 2006). SDM, and to a smaller extent NHMM, are relatively drier, and WRF and CCAM relatively wetter than runoff resulting from empirical scaling. Although WRF has an overall decrease in annual rainfall (Fig. 2), annual runoff is slightly wetter, and this is because the main seasons in which runoff is produced (MAM and JJA) have increases in rainfall, in particular along the high runoff-producing region of the Great Dividing Range in JJA (Fig. 2). This results in an increase in absolute runoff (i.e. mm) in MAM and JJA largely offsetting reductions in runoff in SON and to a lesser extent DJF. Generally, the wetter catchments (in terms of mean annual runoff) have lower absolute change signal (i.e. relatively less wet or dry in percentage terms). Except for CCAM, the two wettest catchments (405205 and 405241) have drying median change signals in the annual runoff. The greater uncertainty from downscaling (at least in terms of percentage changes) occurs in the drier catchments, with the 10–20% reduction in annual runoff projected by empirical downscaling either doubling under SDM to a reversed signal (i.e. 10–20% increase in runoff) under WRF and CCAM.

The seasonal results roughly follow the annual results with a general increase in magnitude of the change signal, consistent with catchment average response to rainfall changes (Chiew, 2006; Potter et al., 2011). Some projected seasonal changes (e.g. CCAM in DJF and WRF in MAM) are for increases over 100%, reflecting the increase in rainfall (Fig. 2). This large percentage change partly reflects the change in the relatively low runoff (percentage change of a small number) in DJF and MAM, and has relatively smaller effect on the annual runoff which is generated mostly in JJA and SON. The most consistent signal in seasonal runoff is for a reduction in SON runoff under all downscaling methods; a response to a consistent change signal for SON rainfall (Fig. 2).

3.4. Projected changes to rainfall and runoff metrics

3.4.1. Dry days and dry spells

The first two rows in Fig. 6 show the projected changes in dry days (number of days with rainfall less than 1 mm) and mean dry spells (mean of consecutive number of days with rainfall less than 1 mm). Empirical scaling scales all the observed historical daily rainfall with

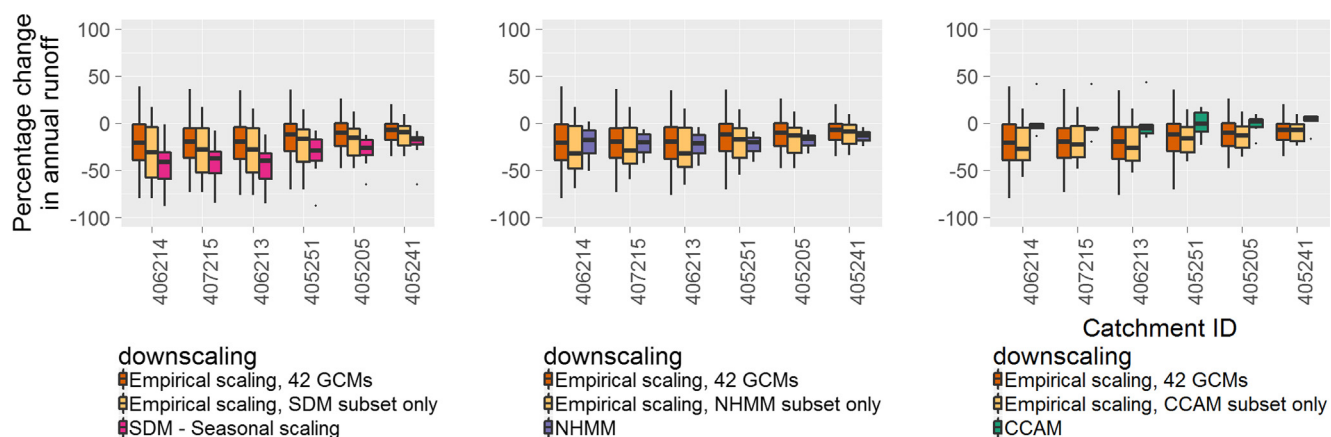


Fig. 4. Effect of GCM selection on empirically scaled change signal compared to downscaled change signals for mean annual runoff. © 2018 CSIRO. All Rights Reserved.

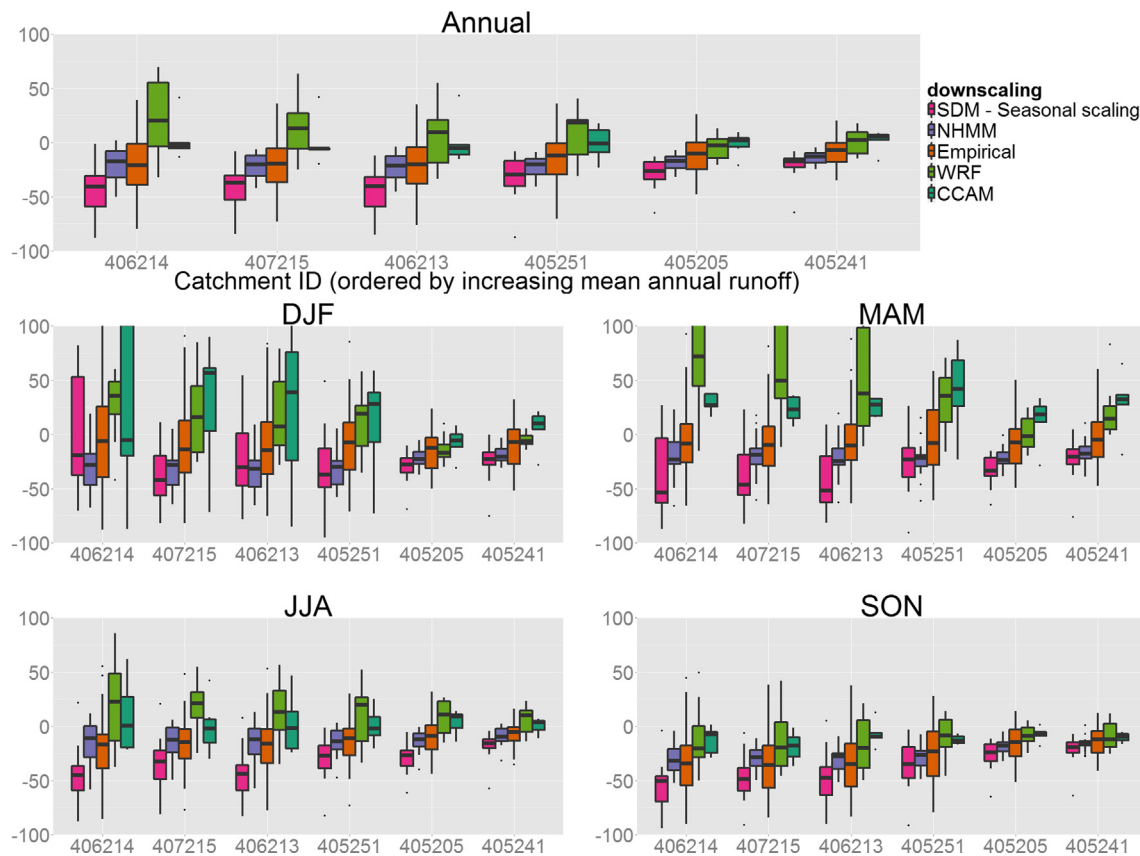


Fig. 5. Distributions of projected annual and seasonal percentage changes to runoff under different downscaling methods. Although some projected seasonal changes (e.g. CCAM in DJF and WRF in MAM) are increases over 100%, the maxima of the scales remain at 100% for ease of comparison. © 2018 CSIRO. All Rights Reserved.

the same scaling factor (for each of the four seasons) informed by the change signal in the GCM, and therefore reflects only a small increase in dry days and longer dry spells consistent with the generally drier future projections in the GCMs. SDM gives similar results, but projects a larger increase in dry days and length of dry spells because of the drier future projections in SDM compared to empirical scaling (or the GCMs). NHMM, which takes into account potential changes in the daily rainfall distribution and sequencing, show considerably larger increases in dry days and length of dry spells compared to empirical scaling and SDM. WRF also shows considerable increases in dry days and length of dry spells despite the relatively smaller projected future mean annual rainfall declines compared to empirical scaling and SDM (see Fig. 2). Despite projections of a wetter future, CCAM results generally show little change in the dry days and dry spells.

3.4.2. High rainfall

High rainfall metrics are represented by the 1st percentile daily rainfall (P1) (third row of Fig. 6) and 3-day maximum rainfall total (fourth row of Fig. 6). Climate change is expected to increase the frequency and intensity of extreme rainfall events over mid-latitudes (IPCC, 2013). Again, the methods that scales the observed historical rainfall (empirical scaling) or use observed historical rainfall analogues (SDM) show decreases in the high rainfall amounts mirroring the projected decline in mean annual rainfall by the GCMs. In the three methods that consider potential changes in the daily rainfall distribution, WRF and CCAM projects higher P1 and 3-day maximum rainfall total, whilst NHMM projects lower P1 and little change in the 3-day maximum rainfall total.

3.4.3. Year-to-year rainfall variability

Fig. 7 shows the coefficient of variability (CV, i.e. the standard deviation divided by the mean) for annual statistics of rainfall. There is

little discernible difference among catchments for all methods. Even more so than the results for changes to dry days and dry spells (Fig. 6), scaling factor approaches do not have noticeable changes in the variability of annual rainfall totals and dry day (less than 1 mm) probabilities. Although annual rainfall can increase or decrease according to the scaling factors employed, the standard deviation changes by a similar order of magnitude, so that CV of annual rainfall is close to zero. Seasonal scaling using SDM and empirical scaling do not change dry day probabilities (Fig. 6) and so dry day variability is likewise unchanged with these methods.

All other methods (NHMM, WRF and CCAM) show a tendency for increased variability of annual rainfall, at least for the ensemble median. The change signal in variability of maximum annual rainfall is less clear with the ensemble ranges covering zero change for most catchments and methods. Although NHMM and WRF project increases in dry day probabilities (Fig. 6), the coefficient of variation in this probability appears to be similar. There is a slight tendency for variability of dry day occurrences to decrease under CCAM.

3.4.4. Low flow

The modelled changes in future low flow metrics (lowest 7-day runoff total, 90th percentile daily runoff (Q90) and number of days with flows less than Q90_H) are presented in the top three rows of Fig. 6. The SDM, NHMM, empirical scaling and WRF results for the three wetter catchments (405251, 405205, 405241) show a smaller 7-day (minimum) low runoff total, lower Q90, and more days with flows below Q90_H in the future. These more severe (or drier) low flow metrics are consistent with the drier average future rainfall and runoff projections. However, there are significant differences in the projections from the different methods, and the range of projections within each method is very large (and much larger than the range of the rainfall characteristics in Fig. 6).

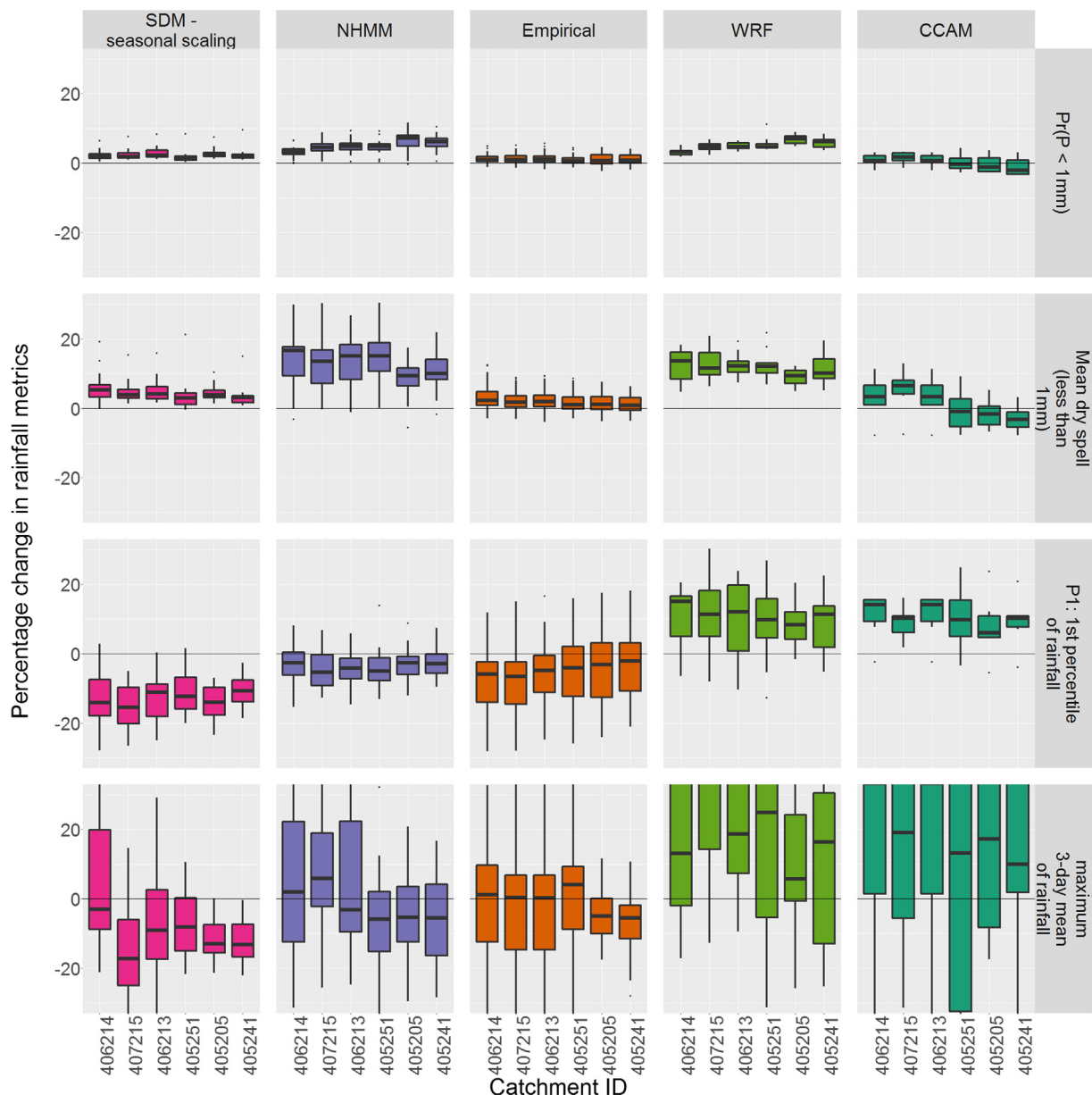


Fig. 6. Percentage changes in selected rainfall metrics (probability of dry day (less than 1 mm) occurrence; mean dry spell (days); 1st percentile of daily rainfall (P1; mm); and 3-day maximum rainfall (mm)) under different downscaling methods. Changes for the last metric extends past -30% to $+30\%$, but the scale has been kept identical across metrics for comparison. © 2018 CSIRO. All Rights Reserved.

Results for the drier catchments are highly variable and difficult to interpret meaningfully, partly because of the difficulty in reliably modelling low flow metrics in dry catchments, and largely because the projected change shown here reflect percentage change to very small numbers (e.g. minimum 7-day runoff total and Q90 are close to zero in these dry catchments). These low flow metrics are dependent on antecedent catchment storage and baseflow recession, the average rainfall amounts, as well as other rainfall-runoff characteristics (and not the dry days in Section 3.4.1 that do not generate runoff) that influence catchment storage and baseflow recession. The rainfall (and runoff) characteristics that influence the low flow metrics can be different for different catchments and hydroclimate and hydrologic regimes, making the low flow metrics difficult to model and predict (Potter et al., 2015, Hughes and Potter, 2015). Rather than providing conclusive change information for the low-flow characteristics in Fig. 8, the results here demonstrate the inherent uncertainty in low-flow metrics, particularly for drier, ephemeral catchments.

3.4.5. High flow

The modelled changes in future high flow metrics (5th percentile daily runoff (Q5), number of days with flows more than Q5_H and 1st percentile daily runoff (Q1)) are presented in the bottom three rows of Fig. 6. These high flow metrics are dependent on changes in the average rainfall (and runoff) as well as the high rainfall events that generate the high flows. The SDM, NHMM and empirical scaling results all show decreases in future high flow metrics, driven by the drier projections of average rainfall and runoff, and decrease in high rainfall (P1) in these methods (see Section 3.4.2 and Fig. 6). WRF shows increases in the high flow metrics, driven by the more intense high rainfalls (see Section 3.4.2 and Fig. 6) and moderated by the small declines in average rainfall (Section 3.1 and Fig. 2). It is interesting to note that despite projections of higher average rainfall and high daily rainfalls in CCAM, there is little change in the modelled future high flow metrics. This may be due to projected changes in rainfall sequencing and serial correlation that are important for high runoff generation yet were not explored in

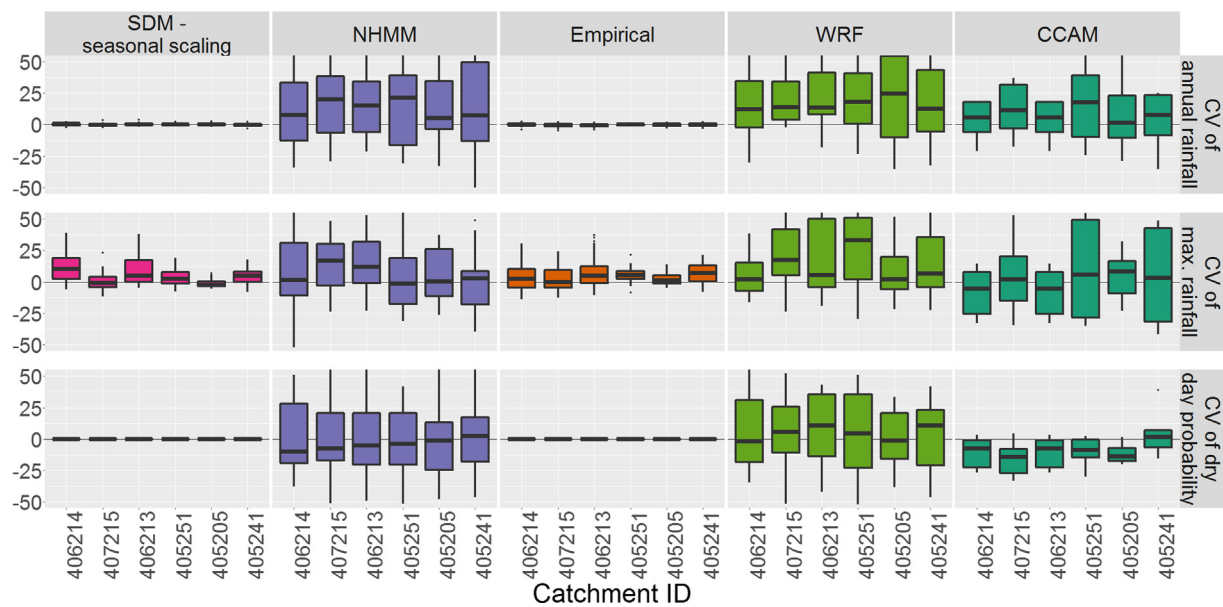


Fig. 7. Percentage changes in selected metrics of interannual rainfall variability: coefficient of variability (CV) of annual rainfall totals; CV of maximum annual rainfall; and CV of dry day (less than 1 mm) occurrence probability. © 2018 CSIRO. All Rights Reserved.

Fig. 6.

3.5. Spatial detail of the downscaling methods

Another potential benefit of downscaling (particularly through dynamic downscaling) is the increased spatial resolution in the change signal. As noted in Table 1, there are significant differences among the spatial resolution of the different downscaling approaches. To illustrate some of the issues and limitations, or indeed potential added value, of different methods, we examine the different spatial detail of the annual rainfall change signal over the six study catchments. Fig. 9 shows mapped rainfall from the five different downscaling methods cropped to the spatial extent of the study catchments. This figure contains the same data from the top row of Fig. 2, although the colour scale limits have been reduced and the number of breakpoints increased to highlight differences within the regional colour scale.

The change signals from empirical scaling are derived directly from the GCM changes. Hence, empirical scaling does not introduce any finer resolution information as demonstrated by the very small spatial variability between and within catchments (Fig. 8 ‘Empirical’). The apparent contour lines in the middle panel of Fig. 2, being ensemble median results, are the result of different GCMs being selected at each grid cell, and can be misleading in terms of spatial signal. The different spatial resolutions of the other downscaling methods (Table 1) are immediately obvious from Fig. 9 with SDM and WRF having the finest resolution change signals ($0.05^\circ \times 0.05^\circ$ and $0.1^\circ \times 0.1^\circ$ respectively) followed by CCAM ($0.5^\circ \times 0.5^\circ$), and NHMM. For NHMM the rainfall is downscaled directly to catchments, and so we expect no within-catchment spatial variability in this method. The between-catchment variability is minimal with NHMM, and this is largely a consequence of the NHMM method, which seeks to preserve spatial correlations in rainfall signals. For CCAM, within-catchment spatial variability in the change signal is present for the larger catchments only, but this resolution is not fine enough to differentiate between the upper and lower catchments in our study domain.

It is only for WRF and SDM that the resolution of change signal is fine enough to provide meaningful within-catchment detail. Especially for the larger catchments, the resolution of WRF appears to be fine enough to be able to distinguish different changes in the upper and lower catchments. For large catchments with distinctive upland and lowland parts, and hence different within-catchment hydrological

behaviour, WRF can potentially provide spatial differences in the change signal, particularly so since the median annual and seasonal change signals (Fig. 2) show evident orographic gradients over the Great Dividing Range, the main geographic feature in Victoria, which corresponds roughly to the ridge of mean annual runoff above 1000 mm in Fig. 1. SDM has the finest scale information, which is commensurate with the underlying gridded climate products commonly in use in Australia (e.g. SILO and AWAP). There is clear spatial variability within even the smallest catchments for this downscaling method. However, Fig. 9 also highlights a potential limitation with this method, which is the sharp divide separating different calibration regions (i.e. the diagonal line from the north of the area to the south-east of the area, see Timbal et al., 2011). SDM selects different climate variables for model calibration in different regions, and this has resulted in variability of the change signal between catchments lying in different regions which is larger than the within-catchment spatial variability. Note however that these spatial differences are much smaller in magnitude than the differences between downscaling methods (cf. Fig. 2). Overall, the finer resolution downscaling methods (i.e. SDM and WRF) appear to show spatial differences in change signal resulting from regional scale topography, however, further assessment is needed to assess the credibility of change signal on these resolutions.

4. Discussion

In a previous comparison of downscaling choices for south-eastern Australia, Frost et al. (2011) recommended using empirical scaling methods for regional water resource planning applications, due principally to the simplicity and robustness of the method. Chiew et al. (2010) found that results from daily scaling (a form of empirical scaling) typically lie within the range of other regionally available downscaling methods, and this is consistent with our results. However, for hydrological applications, empirically downscaled data has two important limitations. Firstly, the scaling factor has the same resolution as the GCM. Thus, change can only be detected on a spatial resolution of around 100–250 km (depending on the resolution of the GCM), which is a limitation for regions where it is reasonable to expect different climate change responses over short spatial ranges, such as along orographic gradients. Secondly, the method cannot capture changes that may influence the sequencing of weather events (as the temporal variability is simply that of the scaled observed data). This is relevant

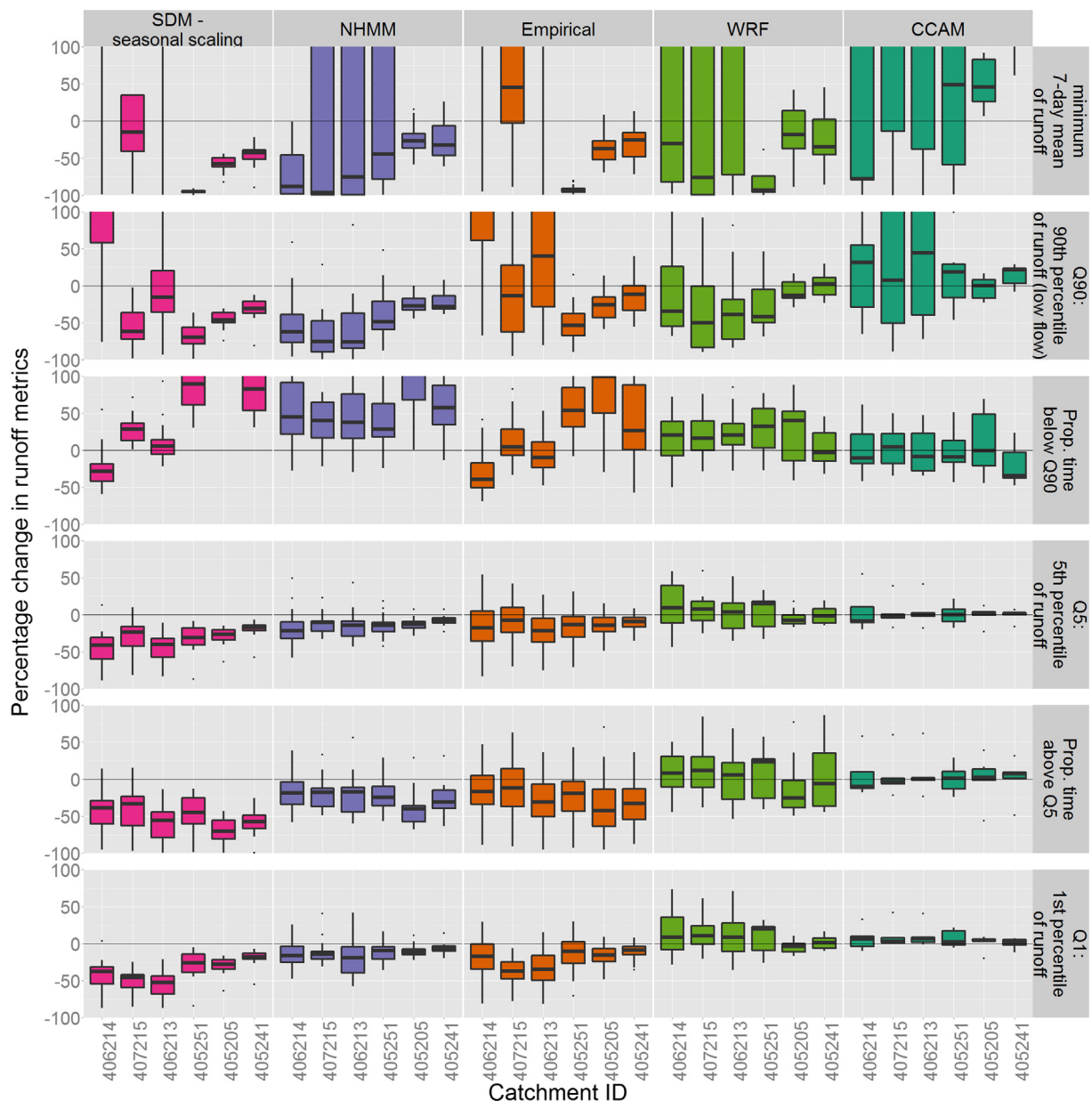


Fig. 8. Percentage changes in runoff metrics (low flow metrics: 7-day minimum flow, 90th percentile of daily flow (Q90), proportion of time spent below historical Q90_H; high flow metrics: 5th percentile of flow (Q5), proportion of time spent above historical Q5_H, and 1st percentile of flow (Q1)) under different downscaling methods. © 2018 CSIRO. All Rights Reserved.

for water-security modelling because catchment runoff relies on large rainfall events in conjunction with longer duration wet spells, and reservoir and water resources management need to cope with multi-year dry sequences. Catchment managers and environmental flow managers also need information about changes to low flow conditions, such as

changes in baseflow and the duration and severity of dry spells impacting water dependent ecosystems and river water quality. The results in section 3.4 demonstrate the challenges in using different downscaled products to provide projections for these metrics. We note that our use of inconsistent (historical) PET time series (as described in

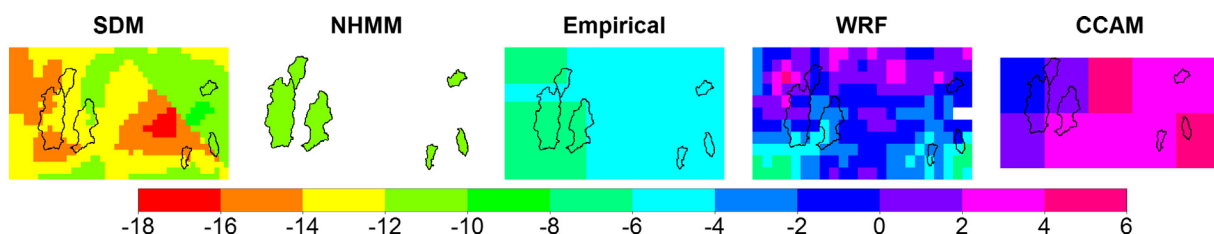


Fig. 9. Differing spatial detail in rainfall change signal (% change in future annual rainfall versus historical annual rainfall). Note the different colour scale compared to Fig. 2 in order to highlight variability within the colour bands of Fig. 2. © 2018 CSIRO. All Rights Reserved.

section 2.2) affects daily runoff prediction more than the annual and seasonal runoff means. However, as noted before, runoff sensitivity to PET is typically an order of magnitude lower than sensitivity to rainfall, and the variability among downscaling methods, particularly for low-flow metrics, is much larger than any expected bias from using historical PET.

We have shown that downscaling can modify the change signal of the GCMs (e.g. Fig. 5). Particularly large changes were associated with runoff based on output from SDM and CCAM. We note that in its full capacity, SDM is a daily resolution downscaling method. As noted in section 2.1.2, we were unable to use the daily data, instead we relied on the monthly data. However, the analysis of regional annual and seasonal rainfall changes using monthly data (Fig. 2) would be identical if we had used daily data. The runoff changes (Fig. 3) are also comparable to recent runoff projections using monthly SDM (Fiddes and Timbal, 2017; Hope et al., 2016). However, projected changes to the seasonal runoff in section 3.3, and particularly the rainfall and runoff metrics in section 3.4, would almost certainly be different using the daily down-scaled data. It is unclear whether WRF modifies the change signal of the GCM as well, as we were not able to assess this against the CMIP5 model ensemble change signals (WRF using GCMs from the CMIP3 archive). In this study we have not examined processes and methodology of the statistical and dynamical downscaling methods. As such, we are unable to state definitively the reason for the wetter or drier change signals in the different downscaling methods, and indeed this may be a feature of the downscaling method in that they are correctly resolving subgrid processes. This is a direction for future research, which would require an in-depth understanding of the biases and sub-grid processes unique to each method.

With respect to sub-setting GCMs based on model skill, recent research (Hope et al., 2016; Grose et al., 2016) has examined the question of whether excluding GCMs that are poor at producing climate characteristics relevant to Victoria's climate. It was found that many of the poorer performing GCMs were underestimating the drying change signal seen for Victoria and that sub-setting the GCM ensemble in this way would produce drier projections. Future research and downscaling assessments could utilise these results to provide a more representative change signal for Victoria and south-eastern Australia. Note however that the GCM selection for SDM, NHMM and CCAM (see section 2.1) was not based on model performance criteria but by and large data availability in the host GCMs.

Climate projections contain uncertainties from many sources, not only downscaling, but uncertainty due to climate scenario, model ensemble, internal variability and rainfall-runoff model and parameterisation (Kay et al., 2009; Chen et al., 2011; Bürger et al., 2013; Ahn et al., 2016). Characterising the contributions of all relevant sources of uncertainty is outside the scope of this paper due to inadequate sample representation of the different modelling stages, but some qualitative assessments can be elucidated from our results. CSIRO and Bureau of Meteorology (2016) illustrated the temporal change in the relative contribution of key sources of uncertainty for regions of Australia following the method described in Hawkins and Sutton (2009). For southern Australia and the time-slice considered here (approximately 50 years into the future) the response uncertainty (model ensemble spread due to response to changing emissions), is larger than scenario uncertainty (ensemble spread due to applied emission scenario), and comparable to uncertainty from internal variability (ensemble spread due to models simulation of internal/natural climate variability). As the projection window extends further into the future, scenario uncertainty becomes relatively more pronounced. Bias correction can potentially be another source of uncertainty, particularly for runoff (Teng et al., 2015). Whereas the relationships between atmospheric variables and precipitation fields in the statistical downscaling models used here are effectively constrained by observed rainfall, rainfall from the dynamical downscaling models are not constrained by observations, necessitating bias correction. Development of bias

correction methods that properly correct precipitation for runoff modelling is a subject of ongoing research.

There are rigorous methods in the literature for the partitioning of uncertainty, such as applying an ANOVA-type approach to decomposing the total variability in change signals into variability from GCM ensemble, climate scenarios and downscaling/bias correction (e.g. Bürger et al., 2013; Woldemeskel et al., 2014; Sunyer et al., 2015). This approach is not applicable to our ensemble because of the different GCM/downscaling combinations (Table 1). However, total variability in, for example, mean annual runoff (Fig. 4) appears to be approximately evenly split between variability from GCM model ensemble and downscaling method. This is comparable to previous studies, which found uncertainty from downscaling is somewhat less than uncertainty from GCM models, but is greater than uncertainty from different rainfall-runoff models (Chiew et al., 2010; Teng et al., 2012b). We conclude that different downscaling methods have significant effects on change signals in rainfall and runoff for southeast Australia, comparable but slightly less than the effect from the GCM ensemble. Scenario uncertainty becomes more important further into the future due to the cumulative impact of increasing emission concentrations in the atmosphere.

5. Conclusions

In the absence of comprehensive downscaling ensembles based on statistical and dynamical downscaling, empirical scaling has been generally recommended, and often used, for regional future water availability assessments in south-eastern Australia. The method is robust, relatively easy to use, computationally less expensive and typically produces responses lying between other downscaling methods. However, the method is restrictive in terms of end-user requirements and expectations who are interested in questions about changes to extreme (high and low) rainfall and runoff events, and seeking information on within-catchment and smaller scale spatial variability of the change signal.

We examined five alternative downscaling methods available for south-eastern Australia at the present time. Consistent with previous downscaling assessments (e.g. Chiew et al., 2010; Frost et al., 2011), the change signals for rainfall and runoff arising from empirical scaling lie between those from the other methods. The statistical downscaling methods (SDM and NHMM) produced drier change signals, and the dynamical downscaling methods (WRF and CCAM) produced wetter change signals in terms of mean annual and seasonal changes to rainfall and runoff. Although there is a large degree of uncertainty, the reductions in rainfall and runoff in SON and to a lesser extent JJA appear to be relatively consistent across the downscaling methods, and this is consistent too with current understanding of recent and future climate trends in the region (Hope et al., 2017).

The results for extreme rainfall and runoff metrics show significantly more differences among the downscaling methods. Although empirical scaling is unable to provide meaningful information on changes to rainfall sequencing (i.e. wet/dry day occurrences), by defining a dry day as one with a non-zero but insignificant amount of rainfall (e.g. 1 mm), changes to dry days and dry spells under empirical scaling appeared to be very similar to those from downscaling methods that do alter the sequencing of rainfall. However, other metrics that rely to a larger extent on the sequencing of rainfall events (e.g. 3-day maximum rainfall) cannot be adequately modelled with empirical scaling. Year-to-year variability in annual rainfall is projected to be slightly larger in the far future time period under non-scaling methods, although uncertainty in this signal remains challenging.

There was reasonable agreement among downscaling methods on the likely changes to low-flow metrics, but only for high-runoff catchments. This is due to low-flow metrics being more difficult to measure and model compared to high and moderate flows, particularly for drier catchments. In this analysis, changes to rainfall extremes, and to a lesser

extent, changes to high-runoff metrics, are largely related to changes to mean rainfall. This gives confidence in the predictability of these metrics under different downscaling methods, noting that large differences in downscaled mean rainfall across south-eastern Australia reduce any confidence level attributed to the projected change in high rainfall and runoff in this region. We also stress that where specific rainfall or runoff metrics are important for an application, we recommend choosing the metrics, models and downscaling methods carefully to be able to make meaningful statements about probable or likely changes to these metrics.

There is potential added value from statistical and dynamical downscaling methods in providing increased spatial information on future change signals. Both SDM and WRF provide output on resolutions that resolve within-catchment variability of rainfall and runoff signals, although further assessment and modelling may be required to fully assess the credibility of change signal on these resolutions.

Acknowledgements

We acknowledge the project funding provided through the Victorian Climate Initiative (VicCI) by the Victorian Department of the Environment, Land, Water and Planning (DELWP). The streamflow data were obtained from the Bureau of Meteorology Water Data Online, and are licensed under Creative Commons Attribution 3.0 Australia (CC BY 3.0 AU) License © State of Victoria (Department of Environment, Land, Water and Planning).

We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP5, and we thank the climate modelling groups (listed in Table 2) for producing and making available their model output. For CMIP5 the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We acknowledge the modeling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP's Working Group on Coupled Modelling (WGCM) for their roles in making available the WCRP CMIP3 multi-model dataset. Support of this dataset is provided by the Office of Science, U.S. Department of Energy.

SDM downscaling data was obtained from the Climate Change in Australia project (<http://www.climatechangeinaustralia.gov.au/>) and we acknowledge Bertrand Timbal and Sonya Fiddes (Bureau of Meteorology) for this data. Assistance with the preparation of the CCAM data was provided by Louise Wilson, Craig Heady and Kim Nguyen (CSIRO). The WRF downscaling data was provided by the NSW Department of Environment & Heritage (<http://climatechange.environment.nsw.gov.au/Climate-projections-for-NSW>) and assistance was provided by David Fuchs (NSW Department of Environment & Heritage) and Jason Evans (UNSW). We thank three anonymous reviewers and the Associate Editor (Professor Ashish Sharma) for their helpful reviews of the manuscript.

References

- Ahn, K.-H., Merwade, V., Ojha, C.S.P., Palmer, R.N., 2016. Quantifying relative uncertainties in the detection and attribution of human-induced climate change on winter streamflow. *J. Hydrol.* 542, 304–1216. <http://dx.doi.org/10.1016/j.jhydrol.2016.09.015>.
- Bürger, G., Sobie, S.R., Cannon, A.J., Werner, A.T., Murdock, T.Q., 2013. Downscaling extremes: an intercomparison of multiple methods for future climate. *J. Clim.* 26, 3429–3449. <http://dx.doi.org/10.1175/JCLI-D-12-00249.1>.
- Charles, S.P., Bates, B.C., Hughes, J.P., 1999. A spatio-temporal model for downscaling precipitation occurrence and amounts. *J. Geophys. Res. Atmos.* 104 (D24), 31657–31669.
- Chen, J., Brissette, F.P., Leconte, R., 2011. Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. *J. Hydrol.* 401, 190–202. <http://dx.doi.org/10.1016/j.jhydrol.2011.02.020>.
- Chiew, F.H.S., McMahon, T.A., 1991. The applicability of Morton's and Penman's evapotranspiration estimates in rainfall-runoff modeling. *J. Am. Water Resour. Assoc.* 27, 611–620. <http://dx.doi.org/10.1111/j.1752-1688.1991.tb01462.x>.
- Chiew, F.H.S., Peel, M.C., Western, A.W., 2002. Application and testing of the simple rainfall-runoff model SIMHYD. In: Singh, V.P., Frevert, D.K. (Eds.), *Mathematical Models of Small Watershed Hydrology and Applications*. Water Resources Publication, Littleton, CO, pp. 335–367.
- Chiew, F.H.S., 2006. Estimation of rainfall elasticity of streamflow in Australia. *Hydrol. Sci. J.* 51, 613–625.
- Chiew, F.H.S., Teng, J., Vaze, J., Post, D.A., Perraud, J.-M., Kirono, D.G.C., Viney, N.R., 2009. Estimating climate change impact on runoff across south-east Australia: method, results and implications of modelling method. *Water Resour. Res.* 45, W10414. <http://dx.doi.org/10.1029/2008WR007338>.
- Chiew, F.H.S., Kirono, D.G.C., Kent, D.M., Frost, A.J., Charles, S.P., Timbal, B., Nguyen, K.C., Fu, G., 2010. Comparison of runoff modelled using rainfall from different downscaling methods for historical and future climates. *J. Hydrol.* 387, 10–23. <http://dx.doi.org/10.1016/j.jhydrol.2010.03.025>.
- Chiew, F.H.S., Zheng, H.X., Potter, N.J., Ekström, M., Grose, M.R., Kirono, D.G.C., Zhang, L., Vaze, J., 2017. Future runoff projections for Australia and science challenges in producing next generation projections. In: Syme, G., Hatton MacDonald, D., Fulton, B. and Piantadosi, J. (eds) *MODSIM2017, 22nd International Congress on Modelling and Simulation*. Modelling and Simulation Society of Australia and New Zealand, December 2017, pp. 1745–1751. <http://www.mssanz.org.au/modsim2017/L16/chiew.pdf>.
- CSIRO, 2008. Water availability in the Murray-Darling Basin. A report to the Australian Government from the CSIRO Murray-Darling Basin Sustainable Yields Project. CSIRO, Australia, pp. 67.
- CSIRO, 2012. South Eastern Australian Climate Initiative Program Annual Report 2011/12. CSIRO, Australia, September 2012, pp. 151.
- CSIRO and Bureau of Meteorology, 2015. Climate change in Australia. Information for Australia's natural resource management regions: Technical report, CSIRO and Bureau of Meteorology, Australia. < <http://www.climatechangeinaustralia.gov.au/en/publications-library/technicalreport/> >.
- DELWP, 2016. Guidelines for assessing the impacts of climate change on water supplies in Victoria. Final version, December 2016. Department of Environment, Land, Water & Planning, available at: <http://delwp.vic.gov.au/water/water-corporations/climate-change-guidelines>. (accessed 02.2017).
- Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.* 28, 1015–1031. <http://dx.doi.org/10.1029/91WR02985>.
- Ekström, M., Grose, M.R., Whetton, P.H., 2015. An appraisal of downscaling methods used in climate change research. *WIREs Climate Change* 6, 301–319. <http://dx.doi.org/10.1002/wcc.339>.
- Ekström, M., Grose, M.R., Heady, C., Turner, S., Teng, J., 2016. The method of producing climate change datasets impacts the resulting policy guidance and chance of mal-adaptation. *Climate Serv. Available Online*. <http://dx.doi.org/10.1016/j.cliser.2016.09.003>.
- Evans, J.P., Ji, F., Abramowitz, G., Ekström, M., 2013. Optimally choosing small ensemble members to produce robust climate simulations. *Environ. Res. Lett.* 8. <http://dx.doi.org/10.1088/1748-9326/8/4/044050>.
- Evans, J.P., Ji, F., Lee, C., Smith, P., Argüeso, D., Fita, L., 2014. Design of a regional climate modelling projection ensemble experiment – NARClIM. *Geosci. Model Dev.* 7, 621–629. <http://dx.doi.org/10.5194/gmd-7-621-2014>.
- Evans, J.P., Argüeso, D., 2014b. Guidance on the use of bias corrected data. NARClIM Technical Note 3, 7pp., NARClIM Consortium, Sydney, Australia.
- Fiddes, S., Timbal, B., 2017. Future impacts of climate change on streamflows across Victoria, Australia: making use of statistical downscaling. *Climate Res.* 71, 219–236. <http://dx.doi.org/10.3354/cr01447>.
- Fowler, H.J., Blenkinsop, S., Tebaldi, C., 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *Int. J. Climatol.* 27, 1547–1578. <http://dx.doi.org/10.1002/joc.1556>.
- Frost, A.J., Charles, S.P., Timbal, B., Chiew, F.H.S., Mehrotra, R., Nguyen, K.C., Chandler, R.E., McGregor, J.L., Fu, G., Kirono, D.G.C., Fernandez, E., Kent, D.M., 2011. A comparison of multi-site daily rainfall downscaling techniques under Australian conditions. *J. Hydrol.* 408, 1–18. <http://dx.doi.org/10.1016/j.jhydrol.2011.06.021>.
- Fu, G., Charles, S.P., Kirchner, S., 2013a. Daily rainfall projections from general circulation models with a downscaling nonhomogeneous hidden Markov model (NHMM) for southeastern Australia. *Hydrol. Process.* 27, 3663–3673.
- Fu, G., Charles, S.P., Chiew, F.H.S., Teng, J., Zheng, H., Frost, A.J., Liu, W., Kirchner, S., 2013b. Modelling runoff with statistically downscaled daily site, gridded and catchment rainfall series. *J. Hydrol.* 492, 254–265. <http://dx.doi.org/10.1016/j.jhydrol.2013.03.041>.
- Grose, M.R., Moise, A.F., Timbal, B., Katzfey, J.J., Ekström, M., Whetton, P., 2015. Climate projections for southern Australian cool-season rainfall: insights from a downscaling comparison. *Clim. Res.* 62 (3), 251–265.
- Grose, M., Risbey, J., Moise, A., Osbrough, S., Heady, C., Wilson, L., Erwin, T., 2016. Constraints on southern Australian rainfall change based on atmospheric circulation in CMIP5 simulations. *J. Climate* 30, 225–242. <http://dx.doi.org/10.1175/JCLI-D-16-0142.1>.
- Harris, R.M.B., Grose, M.R., Lee, G., Bindoff, N.L., Porfiro, L.L., Fox-Hughes, P., 2014. Climate projections for ecologists. *WIREs Clim. Change* 5, 621–637. <http://dx.doi.org/10.1002/wcc.291>.
- Hawkins, E., Sutton, R., 2009. The potential to narrow uncertainty in regional climate predictions. *Bull. Am. Meteorol. Soc.* 90, 1095–1107.
- Hope, P., Timbal, B., Hendon, H., Ekström, M., Day, K. 2016. Victorian Climate Initiative annual report 2015-16. Bureau Research Report BRR-015.
- Hope, P., Timbal, B., Hendon, H., Ekström, M., Potter, N., 2017. A synthesis of findings from the Victorian Climate Initiative (VicCI). Bureau of Meteorology, Australia.

- pp. 55.
- Hughes, J.P., Guttorp, P., Charles, S.P., 1999. A non-homogeneous hidden Markov model for precipitation occurrence. *J. R. Statist. Soc. Ser. C: Appl. Statist.* 48 (1), 15–30.
- Hughes, J.D., Potter, N.J., 2015 (to appear). Investigation into low flow estimation methods in the Mt Lofty Ranges, South Australia, Goyder Institute Technical Report Series No. 15/44. Adelaide, South Australia.
- IPCC (2013) Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. in: Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1535.
- Jones, D.A., Wang, W., Fawcett, R., 2009. High-quality spatial climate data-sets for Australia. *Austral. Meteorol. Oceanogr.* J. 58, 233–248.
- Kay, A.L., Davies, H.N., Bell, V.A., Jones, R.G., 2009. Comparison of uncertainty sources for climate change impacts: flood frequency in England. *Climatic Change* 92, 41–63. <http://dx.doi.org/10.1007/s10584-008-9471-4>.
- McGregor, J.L., 2005. C-CAM: Geometric Aspects and Dynamic Formulation No. 70. CSIRO Atmospheric Research Technical Papers, CSIRO, Australia. < <http://www.cmar.csiro.au/e-print/open/mcgregor.2005a.pdf> > .
- McGregor, J.L., Dix, M.R., 2008. An updated description of the conformal-cubic atmospheric model. In: Hamilton, K., Ohfuchi, W. (Eds.), *High Resolution Simulation of the Atmosphere and Ocean*. Springer, pp. 51–76.
- Meehl, G.A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J.F.B., Stouffer, R.J., Taylor, K.E., 2007. The WCRP CMIP3 multimodel dataset - A new era in climate change research. *Bull. Am. Meteorol. Soc.* 88.
- Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., 2008. Stationarity is dead: Whither water management? *Science* 319, 573–574. <http://dx.doi.org/10.1126/science.1151915>.
- Morton, F.I., 1983. Operational estimates of actual evapotranspiration and their significance to the science and practice of hydrology. *J. Hydrol.* 66, 1–76.
- Nakicenovic, N., Alcamo, J., Davis, G., De Vries, H.J.M., Fenhann, J., Gaffin, S., Gregory, K., Grubler, A., Jung, T.Y., Kram, T., La Rovere, E.L., Michaelis, L., Mori, S., Morita, T., Pepper, W.J., Pitcher, H., Price, L., Riahi, K., Roehrl, A., Rogner, H.H., Sankovski, A., Schlesinger, M., Shukla, P., Smith, S.J., Swart, R.J., Van Rooijen, S., Victor, N., Dadi, Z., 2000. Special Report on Emissions Scenarios. A Special Report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge, Cambridge University Press.
- Petheram, C., Potter, N.J., Vaze, J., Chiew, F.H.S., Zhang, L., 2011. Towards better understanding of changes in the rainfall–runoff relationships during the recent drought in south-eastern Australia. In: Chan, F., Marinova, D. and Anderssen, R.S. (eds) MODSIM2011, 19th International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2011, pp. 3622–3628. <http://www.mssanz.org.au/modsim2011/16/petheram.pdf>.
- Poff, N.L., Brown, C.M., Grantham, T., Matthews, J.H., Palmer, M.A., Spence, C.M., Wilby, R.L., Haasnoot, M., Mendoza, G.F., Dominique, K.C., Baeza, A., 2016. Sustainable water management under future uncertainty with eco-engineering decision scaling. *Nat. Clim. Change* 6, 25–34. <http://dx.doi.org/10.1038/nclimate2765>.
- Potter, N.J., Chiew, F.H.S., 2009. Statistical characterisation and attribution of recent rainfall and runoff in the Murray-Darling Basin. In: Anderssen, R.S. (Ed.), 18th IMACS World Congress – MODSIM09 International Congress on Modelling and Simulation, 13–17 July 2009. Cairns, Australia, pp. 2812–2818.
- Post, D.A., Chiew, F.H.S., Teng, J., Wang, B., Marvanek, S., 2012. Projected changes in climate and runoff for south-eastern Australia under 1°C and 2°C of global warming. A SEACI Phase 2 special report, CSIRO, Australia, pp. 40.
- Potter, N.J., Petheram, C., Zhang, L., 2011. Sensitivity of streamflow to rainfall and temperature in south-eastern Australia during the Millennium drought. In: Chan, F., Marinova, D. and Anderssen, R.S. (eds) MODSIM2011, 19th International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2011, pp. 3636–3642. <http://www.mssanz.org.au/modsim2011/16/potter.pdf>.
- Potter, N.J., Chiew, F.H.S., 2011. An investigation into changes in climate characteristics causing the recent very low runoff in the southern Murray-Darling Basin using rainfall-runoff models. *Water Resour. Res.* 47. <http://dx.doi.org/10.1029/2010WR010333>.
- Potter, N.J., Zhang, L., Petheram, C., Chiew, F.H.S., 2013. Hydrological non-stationarity in southeastern Australia. Climate and Land Surface Changes in Hydrology: proceedings of the IAHS – IAPSO – IASPEI Joint Assembly in Gothenburg, Sweden, July 2013. In: Boegh, E., E. Blyth, D.M. Hannah, H. Hisdal, H. Kunstmann, B. Su, B. and K. K. Yilmaz (eds.) IAHS Publ. 359, pp. 358–363.
- Potter, N.J., Hughes, J.D., Phogat, V., King, D., Bridgman, R., 2015 (to appear). Low flows hydrology in the Mount Lofty Ranges. Final report for Task 3 of Goyder Institute project I.1.4 “Integrated catchment water planning support for Adelaide Mount Lofty Ranges Water Allocation Planning”, Goyder Institute Technical Report Series /28. Adelaide, South Australia.
- Potter, N.J., Chiew, F.H.S., Zheng, H., Ekström, M., Zhang, L., 2016. Hydroclimate projections for Victoria at 2040 and 2065. CSIRO, Australia, doi:10.4225/08/587fad2-c73afa. Available at: <https://publications.csiro.au/rpr/pub?pid=csiro:EP161427>.
- Saft, M., Peel, M.C., Western, A.W., Perraud, J.-M., Zhang, L., 2016. Bias in streamflow projections due to climate-induced shifts in catchment response. *Geophys. Res. Lett.* 43, 1574–1581. <http://dx.doi.org/10.1002/2015GL067326>.
- Skamarock, W.C., Klemp, J.B., 2008. A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *J. Computat. Phys.* 227, 3465–3485. <http://dx.doi.org/10.1016/j.jcp.2007.01.037>.
- Sunyer, M.A., Hündecha, Y., Lawrence, D., Madsen, H., Willems, P., Martinkova, M., Vormoor, K., Bürger, G., Hanel, M., Kriaučiūnienė, J., Loukas, A., Osuch, M., Yücel, I., 2015. Inter-comparison of statistical downscaling methods for projection of extreme precipitation in Europe. *Hydrol. Earth Syst. Sci.* 19, 1827–1847. <http://dx.doi.org/10.5194/hess-19-1827-2015>.
- Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* 93, 485–498. <http://dx.doi.org/10.1175/BAMS-D-11-00094.1>.
- Teng, J., Chiew, F.H.S., Timbal, B., Wang, Y., Vaze, J., Wang, B., 2012a. Assessment of an analogue downscaling method for modelling climate change impacts on runoff. *J. Hydrol.* 472–473, 111–125. <http://dx.doi.org/10.1016/j.jhydrol.2012.09.024>.
- Teng, J., Vaze, J., Chiew, F.H.S., Wang, B., Perraud, J.-M., 2012b. Estimating the relative uncertainties sourced from GCMs and hydrological models in modeling climate change impact on runoff. *J. Hydrometeorol.* 13 (1), 122–139. <http://dx.doi.org/10.1175/JHM-D-11-058.1>.
- Teng, J., Potter, N.J., Chiew, F.H.S., Zhang, L., Wang, B., Vaze, J., Evans, J.P., 2015. How does bias correction of regional climate model precipitation affect modelled runoff? *Hydrol. Earth Syst. Sci.* 19, 711–728. <http://dx.doi.org/10.5194/hess-19-711-2015>.
- Teutschbein, C., Seibert, J., 2010. (2010) Regional climate models for hydrological impact studies at the catchment scale: a review of recent modeling strategies. *Geography Compass* 4, 834–860. <http://dx.doi.org/10.1111/j.1749-8198.2010.00357.x>.
- Timbal, B., Fernandez, E., Li, Z., 2009. Generalization of a statistical downscaling model to provide local climate change projections for Australia. *Environ. Modell. Softw.* 24 (3), 341–358. <http://dx.doi.org/10.1016/j.envsoft.2008.07.007>.
- Timbal, B., Wang, Y., Evans, A., 2011. Downscaling climate change information: an essential ingredient to incorporate uncertainties into adaptation policies. In: Chan, F., Marinova, D., Anderssen, R.S. (Eds.), MODSIM2011, 19th International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, Perth, pp. 1652–1658.
- Timbal, B., Ekström, M., Fiddes, S., Grose, M., Kirono, D., Lim, E.-P., Lucas, C., Wilson, L., 2016. Climate change science and Victoria. Bureau Research Report BRR-014.
- Van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C., Kram, T., Krey, V., Lamarque, J.F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S.J., Rose, S.K., 2011. The representative concentration pathways: an overview. *Climatic Change* 109, 5–31. <http://dx.doi.org/10.1007/s10584-011-0148-z>.
- Viney, N.R., Perraud, J., Vaze, J., Chiew, F.H.S., Post, D.A., Yang, A., 2009. The usefulness of bias constraints in model calibration for regionalisation to ungauged catchments. In: Anderssen, R.S., Braddock, R.D. and Newham, L.T.H. (eds) 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, pp. 3421–3427 < http://www.mssanz.org.au/modsim09/17/viney_I7a.pdf > .
- Woldemeskel, F.M., Sharma, A., Sivakumar, B., Mehrotra, R., 2014. A framework to quantify GCM uncertainties for use in impact assessment studies. *J. Hydrol.* 519, 1453–1465. <http://dx.doi.org/10.1016/j.jhydrol.2014.09.025>.